<table>
<thead>
<tr>
<th>Title</th>
<th>BENFORD'S LAW FOR LINEAR RECURRENCE SEQUENCES (Transcendental Numbers and Related Topics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>NAGASAKA, KENJI; SHIUE, JAU-SHYONG</td>
</tr>
<tr>
<td>Citation</td>
<td>数理解析研究所講究録 (1986), 599: 33-54</td>
</tr>
<tr>
<td>Issue Date</td>
<td>1986-10</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/99590">http://hdl.handle.net/2433/99590</a></td>
</tr>
<tr>
<td>Right</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Textversion</td>
<td>publisher</td>
</tr>
</tbody>
</table>

Kyoto University
BENFORD'S LAW FOR LINEAR RECURRENCE SEQUENCES

BY

KENJI NAGASAKA (長坂 建二)
DEPARTMENT OF MATHEMATICS
FACULTY OF EDUCATION
SHINSYU UNIVERSITY
NISHI-NAGANO, NAGANO
380, JAPAN

AND

JAU-SHYONG SHIUE
DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF NEVADA, LAS VEGAS
4505, MARYLAND PARKWAY
LAS VEGAS, NEVADA
89154, U.S.A.

TRANSCENDENTAL NUMBER THEORY AND ITS RELATED TOPICS
RESEARCH INSTITUTE OF MATHEMATICAL SCIENCES
KYOTO UNIVERSITY, 23.4-25.4 (1986)

1. INTRODUCTION

It has long been known empirically that in most statistical tables expressed in decimal form, the proportion of numbers with the first significant digit less than or equal to $k$ ($k = 1, 2, \ldots, 9$) is approximately $\log_{10}(k+1)$. The base of logarithms is
taken to be 10, unless otherwise stated. We do not admit 0 as a possible first digit, thus the nine first digits, without regard to position of decimal point, do not occur with equal frequency, and many observed tables give a frequency for the occurrence of a given first digit $k$ approximately equal to $\log_{10}(k+1)/k$. Thus the first significant figure is often 1 than any other digit, and the frequency diminishes up to 9.

This peculiar logarithmic distribution of first digits, while not universal, is so common and yet so surprising at first glance that it has given rise to a varied literature, among the authors of which are mathematicians, statisticians, economists, engineers, physicists and amateurs. Simon Newcomb [RR] formulated this law without any actual numerical data. Frank Benford [EE] made many counts from a large body of actual data and numerical tables, owing to which this peculiarity is known as Benford's law after his name.

Several authors, such as W. H. Furry and Henry Hurwitz [II], S. A. Goudsmit and Furry [KK], Roger S. Pinkham [SS] and Ralph A. Raimi [TT] have sought the explanation of this phenomenon by assuming that all physical constants are selected from a population with some scale-invariant underlying distribution and have shown that certain assumptions about this distribution lead to the logarithmic law.

Some writers have pointed out that tables which occur naturally often represent distributions which are mixtures or composite of other distributions, and that the mixing process itself improves the approximation to Benford's law. If $X_i$ are identically distributed independent random variables, consider their product $Y_N = \prod_{n=1}^{N} X_i$. The central limit theorem then shows that the
random variable $\log Y_n \pmod{1}$ approaches uniform distribution as $n$ increases. Thus $Y_n$ tends to obey Benford's law. Furry and Hurwitz [II] stated this result in terms of convolution. A. K. Adhikari and B. P. Sarkar [AA] and Adhikari [BB] proved this result directly for a special distribution and gave a partial converse on scale invariance.

R. W. Hamming [LL] adopts the finite point of view of a computing machine with floating point arithmetic and showed Benford's law.

Thus Benford's law, started as an empirical law, has been studied in various areas, such as probability theory, statistics, computer arithmetics, etc. And Goudsmit and Furry [KK], Furry and Hurwitz [II] and Warren Weaver [XX] recognize "Benford's law is merely the result of our way of writing numbers", or "Benford's law is a built-in characteristic of our number system", or something like that.

With this underlying philosophy, B. J. Flehinger [HH] considered that the smallest population which contains the set of significant figures of all possible physical constants, past, present and future, must be the set of all positive integers. The explanation of Benford's law should, therefore, lie in the properties of the set of integers as represented in a radix number system.

As far as we consider the distribution of the first significant digits, it is quite natural to restrict ourselves to the set of all positive integers. Thus we adopt the set of all positive integers as a model of the population, which would contain significant numbers of all possible numerical constants. From this population, we sample integers according to a certain sampling procedure and determine whether Benford's law holds or not for the resulting sampled integer sequences.
Benford's law was proved to be true for a geometrically sampled integer sequences \( \{2^n\} \), as an application of an ergodic dynamical system in [DD] and also in [UU]. It is clear that the first digit of an integer \( a_n \) is equal to \( k \) if and only if

\[
k \cdot 10^m \leq a_n < (k + 1) \cdot 10^m
\]

for some nonnegative integer \( m \). Then the uniform distribution mod 1 of \( \{\log a_n\} \) is a sufficient condition for Benford's law to hold for \( \{a_n\} \). By making use of this criterion, we treat general geometrical sequences \( \{c \cdot r^n\} \), where \( c \) and \( r \neq 1 \) are positive integers and prove Benford's law except for some special cases.

The definition of uniform distribution mod 1 is as follows:

**DEFINITION.** A sequence \( (x_n) \) of real numbers is said to be uniformly distributed mod 1, if, for every real number \( a \) and \( b \) with \( 0 \leq a < b < 1 \),

\[
\frac{A_N([a, b); (x_n)]}{N} = b - a,
\]

where \( A_N([a, b); (x_n)] \) is the number of indices \( n \) between 1 and \( N \) for which the fractional part \( (x_n) \) falls in \([a, b)\).

Linear recurrence sampling procedures have been considered with a special reference to Fibonacci numbers. R. L. Duncan [GG] proved that the sequence \( \{\log F_n\} \) is uniformly distributed mod 1, where \( F_n \) is the \( n \)-th Fibonacci numbers, which signifies that Benford's law holds for Fibonacci numbers. L. Kuipers [NN] gave another proof of Duncan's result. John L. Brown, Jr. and R. L. Duncan [FF] and L. Kuipers and Jau-Shong Shiue [OO] extended the results in [GG] and [NN] and proved that the sequence \( \{V_n\} \) obeys...
Benford's law, where \( V_n \) satisfies a linear recurrence formula with some restrictions.

K. Nagasaka [PP] succeeded in extending these results and removed unnecessary restrictions on linear recurrence formulae. We pointed out that the condition on the roots of maximum modulus of the characteristic equation for the corresponding linear recurrence formula is essential and indispensable.

Nagasaki and J.-S. Shiue [QQ] give another proof of results in [PP] by using one of van der Corput's difference theorems and extended them to the case that the roots of maximum modulus of the characteristic equation are purely imaginary. In the case that the characteristic equation has two conjugate complex roots of maximum modulus, it is not known whether Benford's law holds for the sequence \( \{V_n\} \), but the asymptotic distribution function mod 1 of the sequence \( \{V_n/V_{n+1}\} \) is represented explicitly by Péter Kiss and Robert F. Tichy [MM], where \( V_n \) satisfies such a second order recurrence.

Another typical integer sequence satisfying Benford's law is the sequence \( \{n!\} \), [JJJ]. Sequences treated above increase rapidly, at least of exponential order. Slowly increasing integer sequences such as prime numbers, polynomials, do not have the limiting frequency distribution for the occurrences of a given first digit \( k \). Thus we need other summation methods.

Flehinger considered successive cumulative averages (Hölder sums) calling their limit the Banach limit. Roy L. Adler and Alan G. Konheim [CC] proved that the Banach limit is a finitely additive measure on the set of all positive integers and assigns zero measure on every finite subset of positive integers. We [PP]
applied the Banach limit to the sequence \( \{P_n\} \) of positive integers generated by a polynomial and proved that Benford's law holds for it in the sense of the Banach limit. The proof needs a long calculation technique on complex analysis.

Peter Schatte named the Banach limit the \( H_\alpha \)-summation method and considered the mantissa distribution in \([VV]\), which is a general notion including Benford's law. He remarked that our result above for polynomial integer sequences \( \{P_n\} \) may be obtained as an application of his theorem 8.3 in \([VV]\) and got a quantitative result as to the speed of convergence \([WW]\).

2. RECURRENCE SEQUENCES OF FIRST ORDER.

In the preceding paper \([4]\), we considered a linear recurrence formula of order 1. The recurrence sequence \( \{h_n\}_{n=1,2,\ldots} \) satisfies the following recursion formula:

\[
(2.1) \quad h_{n+1} = r \cdot h_n + s,
\]

where \( r \neq 1, s \) and \( h_1 \) are positive integers. Then it is proved that Benford's law holds for the sequence \( \{h_n\}_{n=1,2,\ldots} \) except for the case \( r = 10^m \) with \( m \) some nonnegative integer.

In this Section, we consider, instead of (2.1), the following recursion formula of first order:

\[
(2.2) \quad y_{n+1} = r \cdot y_n + f(n), \quad n = 1,2,\ldots,
\]

where \( r \) and \( y_1 \) are positive integers and the range of \( \{f(n)\} \) is also positive integers. Then we obtain

THEOREM 2.1. Let \( \{y_n\}_{n=1,2,\ldots} \) be an integer sequence.
generated by the recursion formula (2.2). If the series
\[ \sum_{n=1}^{\infty} \frac{f(n)}{r^{n-1}} \text{ is convergent, then the sequence } \{ y_n \}_{n=1,2,\ldots} \]
obeyes Benford's law except for the case \( r = 10^m \) with \( m \) some nonnegative integer.

**REMARK 1.** In the case that \( f(n) = s \) for every \( n \), (2.2) is identical to (2.1), so that this Theorem 2.1 contains Theorem 3.2 in [4] as a special case.

In order to prove Theorem 2.1 we need again Lemma 3.1 in [4] and further one of van der Corput's difference theorems in [5], p. 378.

**LEMMA 2.1.** Let \( \{ x_n \}_{n=1,2,\ldots} \) be a sequence of real numbers. If
\[ \lim_{n \to \infty} (x_{n+1} - x_n) = \alpha, \]
where \( \alpha \) is irrational, then the sequence \( \{ x_n \}_{n=1,2,\ldots} \) is uniformly distributed mod 1.

**PROOF.** From the recursion formula (2.2), we have
\[ y_n = r^{n-1}y_1 + r^{n-2}f(1) + \cdots + r f(n-2) + f(n-1) . \]
Let us consider the ratio of consecutive terms of \( \{ y_n \}_{n=1,2,\ldots} : \)
\[ (2.3) \quad y_{n+1} / y_n \]
\[ = \left( \frac{r^n y_1 + r^{n-1}f(1) + r^{n-2}f(2) + \cdots + f(n)}{y_1 + f(1)/r + \cdots + f(n)/r^{n-1}} \right) / \]
\[ \left( \frac{r^{n-1}y_1 + r^{n-2}f(1) + r^{n-3}f(2) + \cdots + f(n-1)}{y_1 + f(1)/r + f(2)/r^2 + \cdots + f(n-1)/r^{n-1}} \right) \]
\[ = \left[ r \cdot y_1 + \left( f(1) + f(2)/r + \cdots + f(n)/r^{n-1} \right) / \right] / \]
\[ \left[ y_1 + \left( f(1) + f(2)/r + \cdots + f(n-1)/r^{n-2} \right) / r \right] . \]
Put
\[ s_{n-1} = f(1) + f(2)/r + \cdots + f(n-1)/r^{n-2}. \]
Then
\[ s_n = f(1) + f(2)/r + \cdots + f(n-1)/r^{n-2} + f(n)/r^{n-1} \]
\[ = s_{n-1} + f(n)/r^{n-1}, \]
which implies that
\[ \lim_{n \to \infty} s_n = \lim_{n \to \infty} s_{n-1} = B > 0, \]
since the sum \( \sum_{n=1}^{\infty} f(n)/r^{n-1} \) is convergent. Taking the limit of (2.3), we get
\[ \lim_{n \to \infty} \left( \frac{y_{n+1}}{y_n} \right) = \left( \frac{r \cdot y_1 + B}{y_1 + B/r} \right) \]
\[ = \left( \frac{r (r \cdot y_1 + B)}{r \cdot y_1 + B} \right) \]
\[ = r. \]
Therefore
\[ \log y_{n+1} - \log y_n \to \log r, \text{ as } n \to \infty. \]
From Lemma 3.1 in [4], \( \log r \) is irrational. Lemma 2.1 asserts that the sequence \( \{ \log y_n \}_{n=1}^{\infty} \) is uniformly distributed mod 1. Reconsidering the same argument as in the first part of the proof of Theorem 3.1 [4], we complete the proof.

\( \text{(Q.E.D.)} \)

NOTE 1. If we don't stick ourselves to positive integer sequences, we can obviously relax assumptions in Theorem 2.1. Indeed, \( r \) may be a positive constant greater than one and not of the form \( 10^m \) for any nonnegative rational number \( m \). \( y_1 \) may also be a given positive rational number and the range of \( \{ f(n) \} \) is nonnegative rational numbers.
3. LINEAR RECURRENCE SEQUENCES OF ORDER 2.

In this Section, we consider a linear recurrence formula \( L(2, a_2, a_1) \) of order 2. The recurrence sequence \( \{ u_n \}_{n=1,2,\ldots} \) satisfies the following recursion formula of order 2:

\[
(3.1) \quad u_{n+2} = a_2 \cdot u_{n+1} + a_1 \cdot u_n, \quad n \geq 1 \quad (a_1 \neq 0),
\]

and its characteristic equation is

\[
(3.2) \quad \lambda^2 = a_2 \cdot \lambda + a_1 \quad (a_1 \neq 0).
\]

**THEOREM 3.1.** If the characteristic equation (3.2) has two real, distinct roots \( \alpha \) and \( \beta \) with \( |\alpha| > |\beta| \) and \( \alpha \) and \( \beta \) are not of the form \( \pm 10^m \) for any nonnegative integer \( m \), then \( \{ u_n \}_{n=1,2,\ldots} \) obeys Benford's law.

**PROOF.** The \( n \)-th term \( u_n \) can be represented by

\[
(3.3) \quad u_n = A \cdot \alpha^{n-1} + B \cdot \beta^{n-1}, \quad n \geq 1,
\]

where \( A \) and \( B \) are constants depending only on \( a_1, a_2, u_1 \) and \( u_2 \). Moreover \( \alpha \cdot \beta \neq 0 \), since \( a_1 \neq 0 \). We then have

\[
\frac{u_{n+1}}{u_n} = \left( \frac{A \cdot \alpha^n + B \cdot \beta^n}{A \cdot \alpha^{n-1} + B \cdot \beta^{n-1}} \right) = \left( \frac{A \cdot \alpha + B \cdot (\beta/\alpha)^{n-1} \cdot \beta}{A + B \cdot (\beta/\alpha)^{n-1}} \right).
\]

Suppose further that \( |\alpha| > |\beta| \) and \( A \neq 0 \), then

\[
\log u_{n+1} - \log u_n = \log \left( \frac{u_{n+1}}{u_n} \right) \to \log \alpha, \quad \text{as} \quad n \to \infty.
\]

\( \log \alpha \) is irrational by Lemma 4.2 in [4] and from Lemma 2.1, the sequence \( \{ \log u_n \}_{n=1,2,\ldots} \) is uniformly distributed mod 1.

In the case that \( A = 0 \), we have

\[
u_n = B \cdot \beta^{n-1}, \quad n \geq 1.
\]

\( \beta \) is not of the form \( \pm 10^m \) for any nonnegative integer \( m \). Then
\[ \log u_{n+1} - \log u_n = \log \left( \frac{u_{n+1}}{u_n} \right) \]
\[ = \log (B \cdot \beta^n) / (B \cdot \beta^{n-1}) \]
\[ = \log \beta, \]

that is irrational. We derive, again from Lemma 2.1 uniform distribution mod 1 for the sequence \( \{ u_n \}_{n=1,2,\ldots} \).

For the case \( |\alpha| = |\beta| \), we may assume, without loss of generality, that \( 0 < \alpha = |\alpha| = |\beta| \), that is \( \beta = -\alpha \), then we can show also by Lemma 2.1 that \( \{ \log u_n \}_{n=1,3,5,\ldots} \) and \( \{ \log u_n \}_{n=2,4,6,\ldots} \) are both uniformly distributed mod 1, from which \( \{ \log u_n \}_{n=1,2,\ldots} \) is uniformly distributed mod 1 too. Hence \( \{ u_n \}_{n=1,2,\ldots} \) obeys Benford's law.

(Q.E.D.)

NOTE 2. This Theorem 3.1 is almost identical to Theorem 4.1 in [4] but we gave another proof using Lemma 2.1, one of the van der Corput's difference theorems. The only difference between this Theorem and Theorem 4.1 in [4] is the additional assumption: \( \beta \) is neither of the form \( \pm 10^m \) for any nonnegative integer \( m \). This assumption is indispensable when \( A = 0 \), but necessary only for the case that \( A = 0 \) in (3.3).

REMARK 2. The condition on \( \alpha \) cannot be removed. Consider \( u_{n+2} = u_n, \ n \geq 1 \), and \( (u_1, u_2) = (c_1, c_2) \), where \( c_1 \) and \( c_2 \) are arbitrary positive integers. The roots of the corresponding characteristic equation are \( \pm 1 = \pm 10^0 \), and the sequence \( \{ u_n \}_{n=1,2,\ldots} \) is purely periodic with period of length 2. Obviously, the sequence \( \{ u_n \}_{n=1,2,\ldots} \) does not obey Benford's law.
THEOREM 3.2. If the characteristic equation (3.2) has a double real root \( \alpha \) which is not of the form \( \pm 10^m \) for any nonnegative integer \( m \), then \( \{ u_n \}_{n=1,2,\ldots} \) obeys Benford's law.

PROOF. We can express the \( n \)-th term \( u_n \) by

\[
u_n = (A \cdot n + B) \cdot \alpha^{n-1}, \quad n \geq 1,
\]

where \( A \) and \( B \) are constants depending only upon \( a_1, a_2, u_1 \) and \( u_2 \). Then

\[
\log u_{n+1} - \log u_n = \log \left( \frac{u_{n+1}}{u_n} \right) = \log \left( \frac{(A \cdot n + A + B) \alpha^n}{(A \cdot n + B) \alpha^{n-1}} \right) = \log |A \cdot n + A + B| / |A \cdot n + B| + \log |\alpha|
\]

\[
\to \log |\alpha|, \quad \text{as } n \to \infty.
\]

Since \( \log |\alpha| \) is irrational, repeating the same argument with Lemma 2.1 as in the proof of Theorem 3.1, we finish the proof.

(Q.E.D.)

REMARK 3. As a general setting throughout, we agree that \( \{ u_n \}_{n=1,2,\ldots} \) is a sequence of positive integers. From the recurrence formula (3.1) with \( \mathbf{a} \) and \( \mathbf{c} \) integral vectors, \( u_n \) may be a negative integer. In this case, we consider the sequence

\[
\{ v_n \}_{n=1,2,\ldots} = \{ |u_n| \}_{n=1,2,\ldots},
\]

instead of \( \{ u_n \}_{n=1,2,\ldots} \) and Theorem 3.1 and Theorem 3.2 holds for the sequence \( \{ v_n \}_{n=1,2,\ldots} \).

REMARK 4. The modulus \( |\alpha| \) in Theorem 3.1 and in Theorem 3.2 is greater than one, since \( \alpha \cdot \beta = a_i \neq 0 \) is an integer. Then \( |u_n| \) tends to infinity as \( n \) tends to infinity possibly except
when \( A = 0 \). Suppose that \( A = 0 \), \( B \neq 0 \) and \( |u_n| \to 0 \) as \( n \to \infty \). Since \( \{ u_n \}_{n=1,2,\ldots} \) is an integer sequence, \( u_n \) is always zero from a certain point on, which is of no interest. If \( A = B = 0 \), this sequence \( \{ u_n \}_{n=1,2,\ldots} \) is the sequence of zeroes, which is of no interest either.

Apart from the characteristic equation (3.2), let us consider a sequence \( \{ u_n \}_{n=1,2,\ldots} \) originally defined by

\[
u_n = C \cdot \gamma^{n-1} + D \cdot \delta^{n-1}, \quad n \geq 1,
\]

where \( C \neq 0, D, \gamma \) and \( \delta \) are real constants and \( 0 < |\delta| < |\gamma| \). Then \( \{ \log |u_n| \}_{n=1,2,\ldots} \) is uniformly distributed mod 1 unless \( \gamma \) is of the form \( \pm 10^m \) with \( m \) a nonnegative rational number. In this situation, \( |\gamma| \) may be smaller than one, i.e. \( u_n \to 0 \), as \( n \to \infty \). Then by considering the first nonzero digits of \( u_n \), Benford's law holds also for \( \{ u_n \}_{n=1,2,\ldots} \). (From Theorem 1 in Persi Diaconis [2].)

If the characteristic equation (3.2) has two complex conjugate roots, \( \alpha \) and \( \overline{\alpha} \), where \( \overline{z} \) is the complex conjugate of \( z \), then the situation is a little confusing. We shall consider only for the case: \( a_2 = 0 \) and \( D = 4a_1 < 0 \).

In this case, two complex conjugate roots are

\[
\alpha = \sqrt{a_1} = a i, \quad \overline{\alpha} = -ai,
\]

by setting \( a = \sqrt{-a_1} \). Then

\[
(3.4) \quad u_n = A(ai)^{n-1} + (-1)^{n-1} \overline{A(ai)}^{n-1}, \quad n \geq 1,
\]

where
A = (c_1 \cdot a - c_2 i) / 2a.

If we suppose further that $A$ is real, then $u_2 = c_2$ must be zero and from (3.1), $u_{2m} = 0$, $m \geq 1$. Original Benford’s law signifies that the distribution of the first significant digits except zero obeys the logarithmic law. Thus we agree to say that Benford’s law holds for an integer sequence $\{a_n\}_{n=1,2,\ldots}$ if the distribution of the first digits of $\{b_n\}_{n=1,2,\ldots}$ obeys the logarithmic law, where $\{b_n\}_{n=1,2,\ldots}$ is the subsequence of all non-zero elements of $\{a_n\}_{n=1,2,\ldots}$.

Direct calculation from (3.4) shows that

$$u_n = 2\text{Re} A \cdot a^{4k}, \quad \text{if } n = 4k + 1,$$

$$= -2\text{Im} A \cdot a^{4k+1}, \quad \text{if } n = 4k + 2,$$

$$= -2\text{Re} A \cdot a^{4k+2}, \quad \text{if } n = 4k + 3,$$

$$= 2\text{Im} A \cdot a^{4k+3}, \quad \text{if } n = 4k + 4.$$

From the above convention, we may suppose, without loss of generality, that $u_n \neq 0$ for arbitrary $n$. Then the following four sequences $\{\log |u_n|\}_{n=1,5,\ldots}$, $\{\log |u_n|\}_{n=2,6,\ldots}$, $\{\log |u_n|\}_{n=3,7,\ldots}$, and $\{\log |u_n|\}_{n=4,8,\ldots}$ are uniformly distributed mod 1 unless $a$ is of the form $10^m$ for some nonnegative integer $m$. Thus $\{|u_n|\}_{n=1,2,\ldots}$ obeys Benford’s law.

Considering Remark 3 and the convention above, we get

**THEOREM 3.3.** If the characteristic equation has two purely imaginary complex roots and $a_1$ is not of the form $-10^m$ for any nonnegative integer $m$, then $\{u_n\}_{n=1,2,\ldots}$ obeys Benford’s law.
4. LINEAR RECURRENCE SEQUENCES OF ARBITRARY ORDER.

In this final Section, we treat a general linear recurrence formula $L(d, \alpha, \xi)$ and the recurrence sequence $\{ u_n \}_{n=1,2,\ldots}$ satisfies the following linear recursion formula of order $d$:

$$u_{n+d} = a_{d-1} u_{n+d-1} + a_{d-2} u_{n+d-2} + \cdots + a_0 u_n, \quad n \geq 1,$$

and also the initial conditions:

$$u_1 = c_1, \quad u_2 = c_2, \quad \cdots \quad \text{and} \quad u_d = c_d,$$

where

$$\alpha = (a_{d-1}, a_{d-2}, \ldots, a_0) \quad \text{and} \quad \xi = (c_1, c_2, \ldots, c_d)$$

are $d$-dimensional integral vectors. The characteristic equation of (4.1) is

$$\lambda^d = a_{d-1} \lambda^{d-1} + a_{d-2} \lambda^{d-2} + \cdots + a_1 \lambda + a_0.$$

Analogously to Theorem 3.2, we get the following Theorem 4.1, which we did not consider in the preceding paper [4].

**THEOREM 4.1.** If the characteristic equation has only one root $\alpha$ of multiplicity $d$ which is not of the form $\pm 10^m$ for any nonnegative integer $m$, then Benford's law holds for the linear recurrence sequence $\{ u_n \}_{n=1,2,\ldots}$.

**PROOF.** By (4.1) and (4.3), we have that

$$u_n = (b_0 + b_1 n + \cdots + b_{d-1} n^{d-1}) \alpha^{n-1},$$

where $b_0, b_1, \cdots$ and $b_{d-1}$ are constants depending only on $\alpha$, $\xi$ and $\alpha$. From Remark 4, we may suppose that $u_n \neq 0$ for any $n$. Then

$$u_{n+1} / u_n = \left( \frac{b_0 + b_1 (n+1) + \cdots + b_{d-1} (n+1)^{d-1}}{b_0 + b_1 n + \cdots + b_{d-1} n^{d-1}} \right) \alpha^{n-1}$$

and

$$u_{n+1} / u_n = \left( \frac{b_0 + b_1 (n+1) + \cdots + b_{d-1} (n+1)^{d-1}}{b_0 + b_1 n + \cdots + b_{d-1} n^{d-1}} \right) \alpha^{n-1}.$$
\[
= \left( b_0 + b_1(n+1) + \cdots + b_d(n+1)^d-1 \right) \cdot \alpha 
\]
\[+ (b_0 + b_1 \cdot n + \cdots + b_d \cdot n^{d-1}) \]\n\[\rightarrow \alpha, \text{ as } n \rightarrow \infty.\]

Thus
\[
\log |u_{n+1}| - \log |u_n| \to \log |\alpha|, \text{ as } n \to \infty.
\]

The number \( \alpha \) is algebraic and therefore \( \log |\alpha| \) is an irrational number. Hence Lemma 2.1 is applicable and we deduce that \( \{ \log |u_n| \}_{n=1,2,\ldots} \) is uniformly distributed mod 1, which indicates that the recurrence sequence \( \{ u_n \}_{n=1,2,\ldots} \) obeys Benford's law.

(Q.E.D.)

Hereafter we suppose that the characteristic equation (4.3) has distinct roots \( \alpha_1, \alpha_2, \ldots \) and \( \alpha_p \) with multiplicity \( m_1, m_2, \ldots \) and \( m_p \), respectively. For our convenience, we arrange the roots \( \alpha_1, \alpha_2, \ldots \) and \( \alpha_p \) according to the magnitude of their modulus, that is,
\[|\alpha_1| \geq |\alpha_2| \geq \cdots \geq |\alpha_p| .\]

It is known that \( u_n \) can be represented by
\[
(4.4) \quad u_n = b_1(n-1) \cdot \alpha_1^{n-1} + b_2(n-1) \cdot \alpha_2^{n-1} + \cdots + b_p(n-1) \cdot \alpha_p^{n-1},
\]
where \( b_1, b_2, \ldots, b_p \) are polynomials of degree at most \( m_1-1, m_2-1, \ldots \) and \( m_p-1 \), respectively. Under this setting, we obtain

**Theorem 4.2.** Suppose that the distinct roots \( \alpha_1, \alpha_2, \ldots \) and \( \alpha_p \) of the characteristic equation (4.3) satisfy
\[(4.5) \quad |\alpha_1| > |\alpha_2| \geq |\alpha_3| \geq \cdots \geq |\alpha_p| \cdot\]

and \(\alpha_1\) is not of the form \(\pm 10^m\) for any nonnegative integer \(m\) and further \(b_1(n-1)\) in (4.4) is not identically zero. Then the linear recurrence sequence \(\{u_n\}_{n=1,2,\ldots}\) obeys Benford's law.

**NOTE 3.** This theorem is identical to Theorem 4.3 in [4]. In order to make clear the situation of the roots, we add an adjective "distinct" and delete the assumption that \(\alpha_1\) is real, since (4.5) indicates that \(\alpha_1\) is real.

Another added assumption on \(b_1(n-1)\) is not so essential. If \(b_j(n-1)\) is the first non-zero polynomial among \(b_1, b_2, \cdots\) and \(b_p\), then (4.5) may be replaced by

\[|\alpha_j| > |\alpha_{j+1}| \geq \cdots \geq |\alpha_p| ,\]

and \(\alpha_j\) is required not being of the form \(\pm 10^m\) for any nonnegative integer \(m\).

**PROOF.** The \(n\)-th term \(u_n\) of the recurrence sequence \(\{u_n\}_{n=1,2,\ldots}\) can be represented by

\[u_n = b_1(n-1) \cdot \alpha_1^{n-1} + b_2(n-1) \cdot \alpha_2^{n-1} + \cdots + b_p(n-1) \cdot \alpha_p^{n-1} ,\]

where \(b_1, b_2, \cdots\) and \(b_p\) are polynomials of degree at most \(m_1-1, m_2-1, \cdots\) and \(m_p-1\), respectively. Considering Remark 4, we may suppose that \(u_n \neq 0\) for any \(n\). Then

\[
\frac{u_{n+1}}{u_n} = \left( \frac{b_1(n) \cdot \alpha_1^n + b_2(n) \cdot \alpha_2^n + \cdots + b_p(n) \cdot \alpha_p^n}{b_1(n-1) \cdot \alpha_1^{n-1} + b_2(n-2) \cdot \alpha_2^{n-1} + \cdots + b_p(n-1) \cdot \alpha_p^{n-1}} \right)
\]
\[= \alpha_1^n \left( b_1(n) + b_2(n) \cdot \frac{\alpha_2}{\alpha_1}^n + \cdots + b_p(n) \cdot \frac{\alpha_p}{\alpha_1}^n \right) / \]

\[\alpha_1^{n-1} \left( b_1(n-1) + b_2(n-1) \cdot \frac{\alpha_2}{\alpha_1}^{n-1} + \cdots + b_p(n-1) \cdot \frac{\alpha_p}{\alpha_1}^{n-1} \right) \]

\[\rightarrow \alpha_1, \text{ as } n \rightarrow \infty.\]

Now

\[\log |u_{n+1}| - \log |u_n| \rightarrow \log |\alpha_1|, \text{ as } n \rightarrow \infty,\]

and \(\log |\alpha_1|\) is irrational. Hence Lemma 2.1 applies and we obtain that \(\{\log |u_n|\}_{n=1,2,\ldots}\) is uniformly distributed mod 1.

This proves \(\{u_n\}_{n=1,2,\ldots}\) obeys Benford's law.

(Q.E.D.)

Now we would like to treat, instead of (4.5), the following case (4.6):

(4.6) \[|\alpha_1| = |\alpha_2| > |\alpha_3| \geq \cdots \geq |\alpha_p|.\]

We suppose further that

(4.7) \[\alpha_2 = -\alpha_1.\]

Then we distinguish two cases;

I. \(\alpha_1\) and \(\alpha_2\) are real:

Hence \(u_n\) can be represented by

\[u_n = \left( b_1(n) + b_2(n) \cdot \frac{\alpha_1}{\alpha_2} + b_3(n) \cdot \frac{\alpha_1}{\alpha_2}^2 + \cdots + b_p(n) \cdot \frac{\alpha_1}{\alpha_2}^{p-1} \right),\]

if \(n\) is odd,

\[= \left( b_1(n-1) - b_2(n-1) \cdot \frac{\alpha_1}{\alpha_2} + b_3(n-1) \cdot \frac{\alpha_1}{\alpha_2}^2 + \cdots + b_p(n-1) \cdot \frac{\alpha_1}{\alpha_2}^{p-1} \right),\]

if \(n\) is even.

Likewise as in Theorem 4.2, we may suppose that \(b_1\) and \(b_2\) are non-zero polynomials and \(b_1 \neq b_2\). Then, for odd \(n = 2m + 1\),
\[ \frac{u_{2m+2}}{u_{2m}} = \frac{\left[ (b_1(2m+1)+b_2(2m+1)) \cdot \alpha_1^{2m+1} + b_3(2m+1) \cdot \alpha_3^{2m+1} + \cdots + b_p(2m+1) \cdot \alpha_p^{2m+1} \right]}{\left[ (b_1(2m-1)+b_2(2m-1)) \cdot \alpha_1^{2m-1} + b_3(2m-1) \cdot \alpha_3^{2m-1} + \cdots + b_p(2m-1) \cdot \alpha_p^{2m-1} \right]} \rightarrow \alpha_1^2, \text{ as } n \to \infty. \]

Thus
\[ \log |u_{2m+2}| - \log |u_{2m}| \to 2 \cdot \log |\alpha_1|, \text{ as } n \to \infty. \]

If $|\alpha_1|$ is not of the form $10^m$ for any nonnegative integer $m$, then \( \{ \log |u_{2m}| \}_{m=1,2,\ldots} \) is uniformly distributed mod 1 and by the same argument \( \{ \log |u_{2m-1}| \}_{m=1,2,\ldots} \) is uniformly distributed mod 1. Thus \( \{ \log |u_n| \}_{n=1,2,\ldots} \) is uniformly distributed mod 1.

II. $\alpha_1$ and $\alpha_2$ are purely imaginary:

In this case, we put
\[ \alpha_1 = ai \quad \text{and} \quad \alpha_2 = -ai, \]
where $a > 0$. Then (4.4) may be rewritten as
\[ u_n = b_1(n-1) \cdot (ai)^{n-1} + b_2(n-1) \cdot (-ai)^{n-1} + b_3(n-1) \cdot \alpha_3^{n-1} + \cdots + b_p(n-1) \cdot \alpha_p^{n-1}. \]

Since
\[ |ai| = |-ai| > |\alpha_3| \geq \cdots \geq |\alpha_p|, \]
and \( \{ u_n \}_{n=1,2,\ldots} \) is a sequence of integers, thus
\[ b_1(n-1) \cdot (ai)^{n-1} + b_2(n-1) \cdot (-ai)^{n-1} \]

is real for sufficiently large every $n$, and consequently
\[ b_1(n-1) \cdot (ai)^{n-1} + b_2(n-1) \cdot (-ai)^{n-1} \]
\[ = b_1(n-1) \cdot (-ai)^{n-1} + b_2(n-1) \cdot (ai)^{n-1} \]
for every \( n \). Thus we get
\[ b_1(n-1) = b_2(n-1), \] for every \( n \).

As we have seen before, the distribution of \( \{ \log |u_n| \}_{n=1,2,\ldots} \) depends only upon \( (n-1) \cdot \log |\alpha_1| \). Analogously to the proof of Theorem 3.3, we consider four subsequences of \( \{ \log |u_n| \}_{n=1,2,\ldots} \) and if \( a \) is not of the form \( 10^m \) for any nonnegative integer \( m \), then each of four subsequence of \( \{ \log |u_n| \}_{n=1,2,\ldots} \) is uniformly distributed mod 1.

Hence \( \{ \log |u_n| \}_{n=1,2,\ldots} \) is uniformly distributed mod 1 and the original sequence \( \{ u_n \}_{n=1,2,\ldots} \) obeys Benford’s law, using the convention in the last Section, if necessary. Thus we get

**THEOREM 4.3.** Suppose that the distinct roots \( \alpha_1, \alpha_2, \ldots \) and \( \alpha_\rho \) of the characteristic equation (4.3) satisfy (4.6) and (4.7) and \( \alpha_1 \) is not of the form \( \pm 10^m \) for any nonnegative integer \( m \) and further \( b_1(n-1) \) and \( b_2(n-1) \) in (4.4) are neither identically zero nor identically equal mutually. Then the linear recurrence sequence \( \{ u_n \}_{n=1,2,\ldots} \) obeys Benford’s law.

**REMARK 5.** We fix the base of logarithms to be 10, but if we change the base to an arbitrary positive integer \( g > 1 \), our arguments still remain valid by exchanging the assumption on \( \alpha_1 \) from \( 10^m \) to \( g^m \).
REFERENCES


